The EMBO Journal vol.15 no.13 pp.3385–3393, 1996

Engrailed and Hox homeodomain proteins contain a related Pbx interaction motif that recognizes a common structure present in Pbx

Lucy T.C.Peltenburg[1] and Cornelis Murre[2]

Department of Biology 0366, University of California, San Diego, La Jolla, CA 92093, USA

[1]Present address: Department of Clinical Oncology, University Hospital, Leiden, The Netherlands

[2]Corresponding author

Hox gene products have the ability to interact with either extradenticle or pbx gene products to bind cooperatively to DNA. The region in Hox proteins that is required for this interaction is located N-terminal of the homeodomain and contains a highly conserved hexapeptide. We now show that the engrailed gene products also contain a Pbx interaction motif positioned within a previously conserved region, the EH2 domain. The EH2 domain is located N-terminal of the homeodomain. Two tryptophan residues present in the *Drosophila* and murine Engrailed EH2 domain are required for cooperativity with extradenticle and Pbx, respectively. A second conserved domain, EH3, is required as well for cooperativity with Pbx, since deletions or an insertion in this region reduce cooperative DNA binding. Peptides containing the Pbx interaction motif of either Engrailed or Hox are capable of destabilizing Engrailed–Pbx and Hox–Pbx cooperative DNA binding. These data indicate that the Pbx interaction motifs present in Hox and engrailed gene products recognize a common structure present in the Pbx family of homeodomain proteins.

Keywords: cooperative DNA binding/Engrailed/homeodomain proteins/Hox/Pbx

Introduction

Homeotic selector proteins are part of a vast transcriptional network involved in the formation of body structures during embryogenesis in a wide variety of organisms (McGinnis and Krumlauf, 1992; Krumlauf, 1994). The homeotic proteins all contain a 60 amino acid region called a homeodomain, which is formed by a flexible N-terminal arm followed by three α-helices and directs binding to A/T-rich DNA sequences (Laughon and Scott, 1984; Desplan *et al.*, 1988). During development of *Drosphila melanogaster*, the characteristic morphology of each of the body segments is dictated by expression of the genes of the HOM cluster along the anterior–posterior axis (McGinnis and Krumlauf, 1992). In vertebrates, the *hox* genes are located in four clusters in the same order as they are expressed along the anterior–posterior axis of the developing organism (Duboule and Dollé, 1989; Wilkinson *et al.*, 1989; Krumlauf, 1994; Lawrence and Morata, 1994).

Although homeodomain proteins have distinct *in vivo* functions, they bind to similar DNA sequences *in vitro* (Desplan *et al.*, 1988; Hoey and Levine, 1988). To account for the biological specificity, it has been proposed that other proteins enhance the DNA binding specificity (Kuziora and McGinnis, 1989; Gibson *et al.*, 1990; Lin and McGinnis, 1992; Chan and Mann, 1993; Zeng *et al.*, 1993). One such co-factor, designated extradenticle (exd), has been identified recently (Flegel *et al.*, 1993; Rauskolb *et al.*, 1993). Through genetic analysis, exd has been shown to act in concert with other homeotic proteins to regulate common downstream target genes, for example *decapentaplegic* (*dpp*) (Peifer and Wieschaus, 1990). *dpp* is positively regulated by both exd and the homeotic gene product Ubx (Rauskolb and Wieschaus, 1994). In addition to acting in parallel with homeotic proteins, exd modulates the activity of the homeodomain protein Engrailed (Peifer and Wieschaus, 1990).

The genetic data describing a synergy between the homeotic gene products and exd recently have been supported by *in vitro* binding assays, which provided a biochemical rationale: cooperative DNA binding. Using binding sites derived from the *dpp* enhancer as well as synthetic templates, exd was demonstrated to have the ability to interact cooperatively with Ubx, abd-A and Engrailed (Chan *et al.*, 1994; van Dijk and Murre, 1994).

The studies using *Drosophila* exd proteins have been extended employing mammalian homologs. Three highly conserved homologs, designated pbx1, pbx2 and pbx3, have been identified in the murine and human genome (Monica *et al.*, 1991). pbx1 was first characterized from the t(1;19) translocation product frequently found in pediatric pre B-cell leukemia (Kamps *et al.*, 1990; Nourse *et al.*, 1990). pbx2 and pbx3 are closely related to pbx1 in 90% of their coding sequence. All three proteins are widely expressed and contain a rather divergent homeodomain (Monica *et al.*, 1991). Not only are the amino acid sequences of the Pbx proteins closely related to exd, but they also have the ability to modulate the binding activity of the mammalian Hox and engrailed gene products to synthetic binding sites (Chang *et al.*, 1995; Knoepfler and Kamps, 1995; Lu *et al.*, 1995; Phelan *et al.*, 1995; van Dijk *et al.*, 1995). Furthermore, Pbx proteins have been shown to interact with Hoxb-1 to bind cooperatively to binding sites present in the Hoxb-1 promoter (Pöpperl *et al.*, 1995).

The domains required for interaction of several Hox proteins with Pbx have been identified recently. A highly conserved hexapeptide sequence present in Hox proteins, separated by a linker region from the N-terminus of the homeodomain, is required for cooperative DNA binding with Pbx1 and Pbx2 (Chang *et al.*, 1995; Johnson *et al.*, 1995; Knoepfler and Kamps, 1995; Neuteboom *et al.*, 1995; Phelan *et al.*, 1995). In addition, the homeodomain

seems to be required as well, since fusions to other homeodomains are not functional in cooperative binding (Neuteboom *et al.*, 1995; Peers *et al.*, 1995). Whereas the molecular requirements of Hox–Pbx interactions are relatively well understood, the interaction domains present in Engrailed have not yet been identified. We show here that a conserved region located in close proximity to the Engrailed homeodomain, designated EH2, is crucial for cooperative DNA binding with Pbx and exd. This region is related to the hexapeptide present in Hox proteins in that it shares a number of residues: proline, tyrosine and tryptophan. Amino acid substitutions of the tryptophans in mammalian and *Drosophila* Engrailed abolish cooperativity with both Pbx and exd. A highly conserved linker region, separating the EH2 domain from the homeodomain, is important as well.

Interestingly, a synthetic peptide centered by the conserved tryptophans is capable of destabilizing both Engrailed–Pbx and Hox–Pbx cooperative binding. Similarly, a Hoxb-7 synthetic peptide that includes the hexapeptide has the ability to abolish these cooperative heterocomplexes. These data support a model in which Hox and Engrailed proteins use overlapping interaction surfaces on Pbx in order to bind DNA cooperatively.

Results

The N- and C-termini of En-2 are not required for cooperative DNA binding with Pbx

Recently, we showed that the *Drosophila* protein exd has the ability to modulate the DNA binding activity of the Engrailed protein, en (van Dijk and Murre, 1994). Their mammalian homologs Pbx and En-2 are also capable of cooperative DNA binding (van Dijk *et al.*, 1995), indicating that the interaction domains involved have been conserved during evolution. Analysis of the cross-species conservation among *En* genes of vertebates and invertebrates displayed five regions of high amino acid conservation, designated EH1–EH5, most of which are surrounding the Engrailed homeodomain (Logan *et al.*, 1992) (Figure 1A).

To study the involvement of the homeodomain-flanking regions EH1, EH2, EH3 and EH5 in cooperative DNA binding with Pbx family members, we constructed combined N- and C-terminal deletions of the murine En-2 coding region (Figure 1B). These homeodomain-containing products were *in vitro* transcribed and translated using reticulocyte lysate and analyzed by electrophoretic mobility shift assays (EMSA). To examine cooperative interactions, we used a probe that was shown previously to allow cooperative binding of En-2–Pbx, 5′-GTCAATTAAATGATCAATCAATTTCG-3′ (van Dijk *et al.*, 1995). Whereas DNA binding by full-length En-2 by itself is almost undetectable, its truncated versions lacking the N- and C-terminal regions show increased DNA binding, possibly due to the relief of a masking effect of one or both terminal domains (Figure 1B, lanes 2–4).

Since cooperative complexes might be obscured by bands generated by the unprogramed reticulocyte lysate itself (Figure 1B, lane 1), we made use of Flag-tagged versions of the Pbx1, Pbx2 and Pbx3 proteins. By incubating with an anti-Flag antibody, the Pbx-containing supershifted complexes can be discriminated more easily from

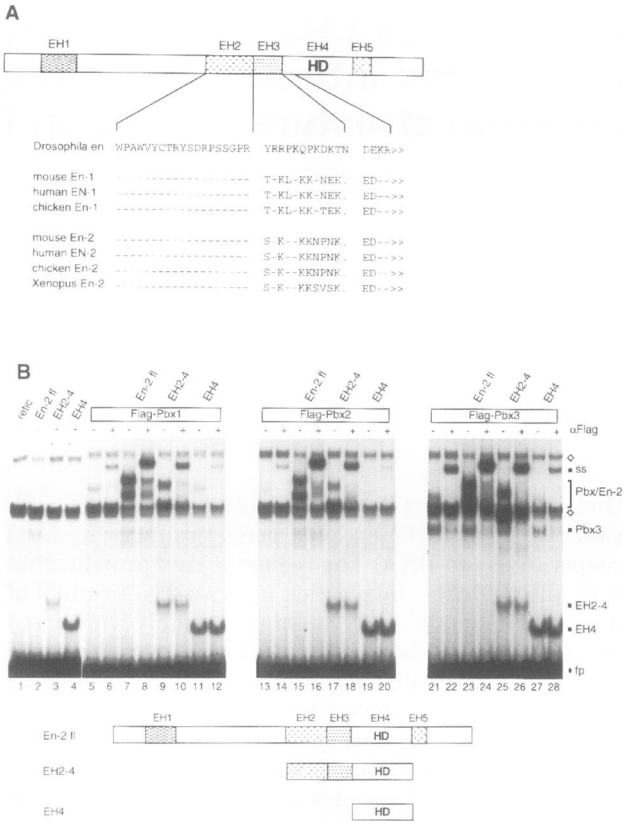

Fig. 1. En-2 conserved domains EH2–4 are sufficient for cooperative DNA binding with Pbx proteins. (**A**) Schematic representation of the Engrailed protein. Approximate locations of the conserved domains EH1, 2, 3, 4 and 5 are indicated. Sequences of EH2, EH3 and the four N-terminal residues of the homeodomain (EH4) of some Engrailed proteins are given below (Logan *et al.*, 1992; Bürglin, 1994). Dashes depict residues identical to the corresponding residue in *Drosophila* en, a dot represents absence of the residue. (**B**) Two μl of the indicated translated proteins were pre-incubated with probe 5′-GTCAATTAAATGATCAATCAATTTCG-3′ and, if applicable, incubated with 1 μl of anti-Flag antibody (αFlag). A schematic representation of the En-2 truncations used is given below. The results were analyzed by EMSA. The lane labeled 'retic' represents incubation of 4 μl of the unprogramed reticulocyte lysate only. On the right, specific complexes are indicated with a dark square or bracket; ss indicates the supershifted complexes; fp the free probe, and complexes formed by unprogramed lysate are indicated with an open diamond.

the background reactivity. In the absence of En-2, weak DNA binding of Pbx1 and Pbx2 (Figure 1B, lanes 5, 6, 13 and 14) and stronger binding of Pbx3 (Figure 1B, lanes 21 and 22) can be detected. In the presence of full-length En-2, the DNA binding activity is increased and new complexes with slower mobility are formed (Figure 1B, lanes 7, 8, 15, 16, 23 and 24). These data demonstrate that Pbx3, like its family members Pbx1 and Pbx2 and exd, shows cooperativity with Engrailed proteins. Deletion of the En-2 N- and C-terminal regions, encompassing EH1 and EH5, does not affect cooperative binding with all three Pbx proteins (Figure 1B, lanes 9, 10, 17, 18, 25 and 26), indicating that EH1 and EH5 are not required for cooperative DNA binding. Removal of the region containing EH2 and the homeodomain-flanking region EH3 leads to a complete loss of cooperative DNA binding (Figure 1B, lanes 11, 12, 19, 20, 27 and 28). These results indicate that the region containing EH2 and EH3 is

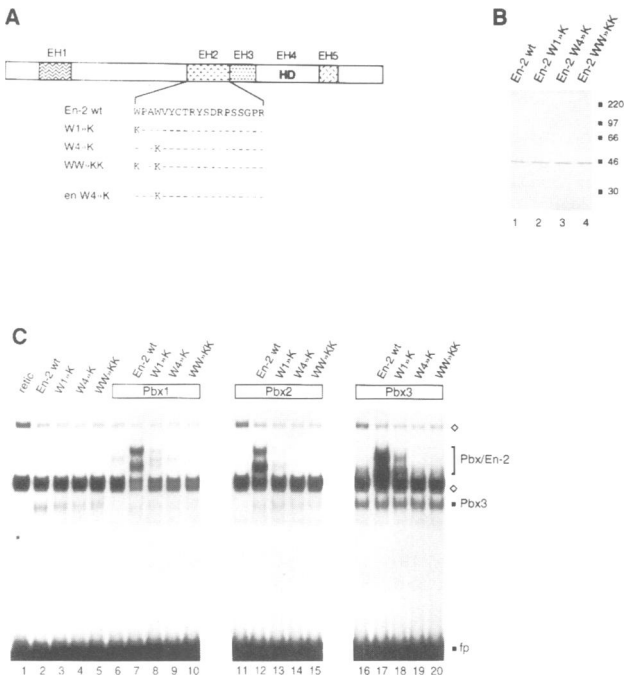

Fig. 2. Mutations in the N-terminal region of EH2 abrogate cooperative binding with Pbx proteins. (**A**) Schematic representation of the mutations generated in full-length murine En-2 and *Drosophila* en. The sequence of EH2 only is given. Dashes indicate residues identical to the wild-type (wt) sequence. (**B**) SDS–gel electrophoresis of [^{35}S]methionine-labeled representative samples of *in vitro* translation reactions of wt and mutant murine En-2 proteins, as indicated. Migration of a protein size marker is depicted in kilodaltons on the right. (**C**) Two μl of each of the indicated *in vitro* translated proteins were incubated with probe 5′-GTCAATTAAATGATCAATCAATT-TCG-3′ and analyzed by EMSA. The lane marked 'retic' represents incubation of 4 μl of the unprogramed reticulocyte lysate only. Specific complexes are indicated on the right with a dark square or bracket; complexes formed by unprogramed lysate are indicated with an open diamond and fp indicates the free probe.

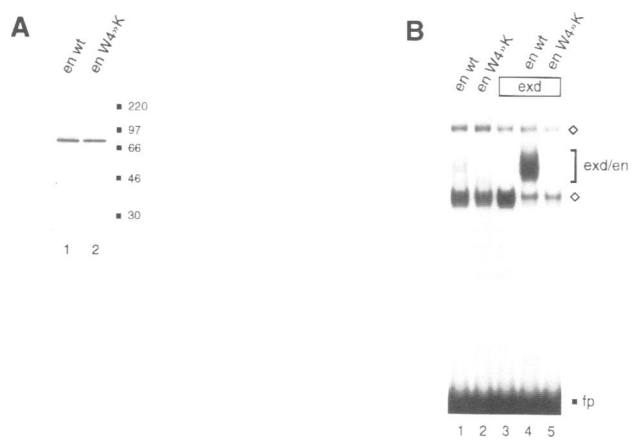

Fig. 3. The requirement for the EH2 domain is conserved in the *Drosophila* Engrailed protein. (**A**) SDS–gel electrophoresis of [^{35}S]methionine-labeled representative samples of *in vitro* translation reactions of wt and mutant *Drosophila* en proteins, as indicated. Migration of a protein size marker is depicted in kilodaltons on the right. (**B**) Two μl of each of the indicated *in vitro* translated proteins were incubated with probe 5′-GTCAATTAAATGATCAATCAATT-TCG-3′ and analyzed by EMSA. Dark squares and brackets indicate specific complexes; fp indicates free probe and the open diamonds mark complexes formed by unprogramed lysate.

indispensable for cooperative interactions involving Pbx and En-2.

The N-terminal region of EH2 is required for cooperativity with Pbx

We and others demonstrated previously that several Hox homeobox proteins bind cooperatively with Pbx proteins to DNA (Chang *et al.*, 1995; Lu *et al.*, 1995; Phelan *et al.*, 1995; van Dijk *et al.*, 1995). In Hox proteins, a highly conserved hexapeptide and the region connecting this motif to the homeodomain are required for cooperative interactions with Pbx1 and Pbx2 (Chang *et al.*, 1995; Johnson *et al.*, 1995; Knoepfler and Kamps, 1995; Neuteboom *et al.*, 1995; Phelan *et al.*, 1995). Comparisons of the hexapeptide sequences of several Hox proteins capable of cooperative binding with Pbx reveal some amino acid similarities with the N-terminal portion of Engrailed EH2. It contains proline, tryptophan and tyrosine residues. Mutation analysis of Hoxb-8 indicated that, of these, the tryptophan and tyrosine are key to cooperative interaction (Neuteboom *et al.*, 1995). To study whether the two Engrailed tryptophans are required, we mutated in En-2 either one or both of these residues to lysines (Figure 2A). The *in vitro* transcribed and translated products were analyzed by SDS–gel electrophoresis (Figure 2B). Wild-type and mutant proteins were translated with similar efficiency and examined for their ability to cooperate with Pbx proteins in DNA binding by EMSA. As for the wild-type En-2 protein, DNA binding of the proteins in the absence of Pbx is hardly detectable (Figure 2C, lanes 2–5). Cooperativity with all three Pbx proteins is affected greatly by mutation of the first tryptophan residue of EH2 (Figure 2C, lanes 8, 13 and 18). Mutation of the second tryptophan or both tryptophan residues leads to a complete loss of En-2–Pbx complexes (Figure 2C, lanes 9, 10, 14, 15, 19 and 20). These data show that the N-terminal region of the conserved EH2 domain is essential for cooperative DNA binding of En-2 with Pbx and that, as in Hox gene products, tryptophan residues play a crucial role.

Evolutionary conserved residues are involved in cooperativity of extradenticle and Engrailed

The amino acid sequence of murine EH2 is completely identical to the EH2 domain of *Drosophila* Engrailed (Logan *et al.*, 1992) (Figure 1A). Previously, we showed that the *Drosophila* homologs of En-2 and Pbx, en and exd, are capable of cooperative DNA binding (van Dijk and Murre, 1994). To determine whether similar residues are involved in interactions of the mammalian and *Drosophila* proteins, the second tryptophan in the *Drosophila* EH2 domain was mutated to a lysine (Figure 2A). The translated products were analyzed by SDS–gel electrophoresis and subjected to EMSA in the presence and absence of *in vitro* translated exd. Wild-type and mutant proteins were translated with approximately the same efficiency (Figure 3A). Whereas DNA binding by either en or exd is undetectable (Figure 3B, lanes 1–3), a strong cooperative complex is seen when wild-type en and exd are incubated together (Figure 3B, lane 4). The mutated en product is completely devoid of cooperative DNA binding (Figure 3B, lane 5), indicating that *Drosophila* and mammalian proteins use identical residues to mediate Engrailed–Pbx interactions.

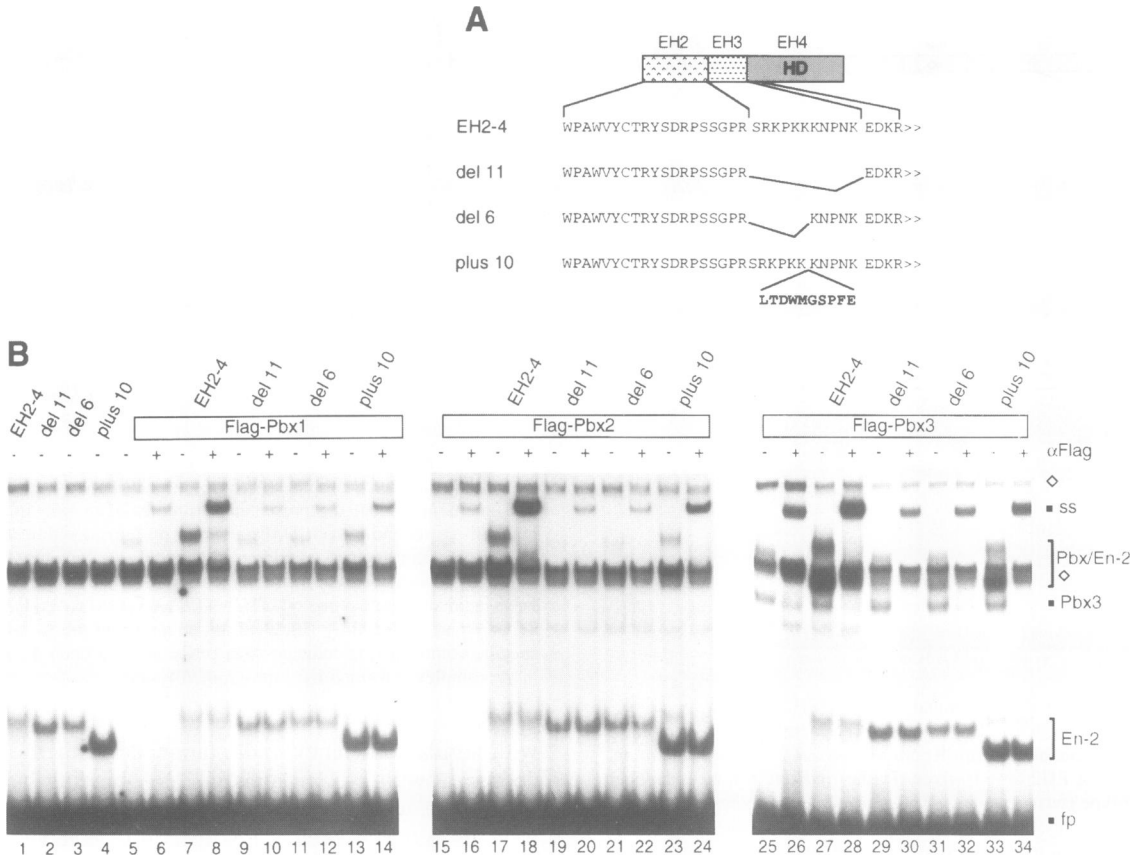

Fig. 4. Conserved domain EH3 is required for optimal cooperativity of Pbx and En-2 DNA binding. (**A**) Schematic representation of the used deletion and fusion proteins. The sequences of domain EH2, EH3 and the four N-terminal amino acids of EH4 are given. The translated protein del 11 lacks region EH3 completely, in del 6 the N-terminal six amino acids of EH3 have been removed. The plus 10 protein has an insertion of 10 residues derived from the linker peptide of the *Drosophila* abd-A protein between the sixth and seventh amino acid of EH3. (**B**) Two μl of the *in vitro* translated proteins were incubated with probe 5′-GTCAATTAAATGATCAATCAATTTCG-3′ in the presence or absence of 1 μl of anti-Flag antibody, as indicated, and analyzed by EMSA. Specific complexes are indicated on the right with a dark square or bracket; ss indicates the supershifted complexes, fp free probe, and complexes formed by unprogramed lysate are indicated with an open diamond.

The EH3 domain is required for interaction with Pbx

Previously, we demonstrated that the linker region present in Hox proteins is important for interaction with Pbx (Neuteboom *et al.*, 1995). The location of the Hox hexapeptide with respect to the homeodomain is similar to the location of the EH2 domain in Engrailed proteins. To establish a role for the EH3 domain as a linker region, we engineered coding regions lacking the N-terminal six residues or all 11 amino acids of domain EH3, as well as a product with an insertion of 10 amino acids derived from the linker peptide of the *Drosophila* abd-A protein (Figure 4A). *In vitro* transcribed and translated proteins were incubated in the presence and absence of Flag-tagged Pbx1, Pbx2 or Pbx3 and analyzed by EMSA. In the absence of Pbx, the truncated En-2 proteins show considerable DNA binding activity (Figure 4B, lanes 1–4). Incubation of EH2–4 with either Pbx1, 2 or 3 again results in complexes with slower mobility (Figure 4B, lanes 7, 8, 17, 18, 27 and 28), but co-incubations of Pbx proteins with the deletion products del 11 or del 6 do not show any cooperative complexes (Figure 4B, lanes 9–12, 19–22 and 29–32). Insertion of 11 residues in the middle of the EH3 domain affects the cooperative binding of Pbx and EH2–4, but some cooperative binding is still observed (Figure 4B, lanes 13, 14, 23, 24, 33 and 34). These results indicate that the EH3 domain is required for cooperative binding of Pbx and En-2, and suggest that the length of the amino acid stretch between EH2 and the homeodomain is critical for optimal cooperativity.

En-2 and Hox peptides spanning the Pbx interaction domains compete for cooperative complexes

The results described above indicate that a tryptophan-containing domain present in EH2 is required for cooperative DNA binding with Pbx. A related motif present in Hox proteins, called the hexapeptide, previously has been shown to be necessary for interaction with Pbx. These data raise the question of whether the tryptophans in EH2 and in the Hox hexapeptide recognize identical residues in Pbx. To resolve this issue, we synthesized a 23 amino acid En-2 peptide, centered by the second tryptophan (TLGAQPMLWPAWVYCTRYSDRPS, Figure 5A), and a corresponding mutant peptide (TLGAQPM-LKPAKVYCTRYSDRPS).

A mixture of *in vitro* translated full-length En-2 and Pbx1 was incubated with increasing concentrations (5–20 μM) of the peptides to allow competition of cooperative complexes. After the addition of the labeled DNA probe containing Pbx and Engrailed recognition sites, the reactions were analyzed by EMSA. In the absence of peptide,

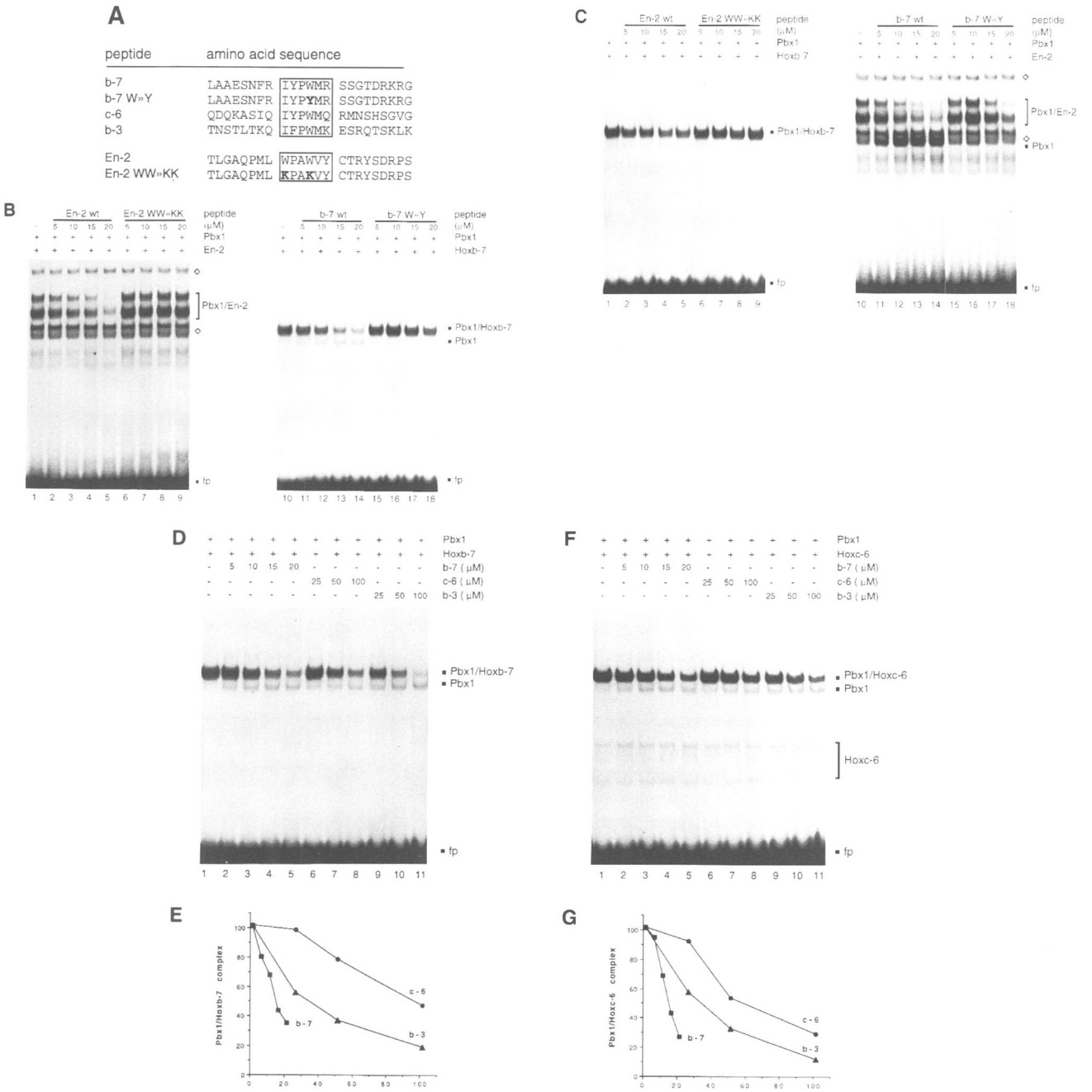

Fig. 5. Peptides spanning the En-2 and Hox interaction domains can cross-compete cooperative Pbx complexes. (**A**) Amino acid sequences of the synthesized wt and mutant peptides, as indicated. The Hox hexapeptide sequences are indicated by the lower box and the corresponding En-2 region by the upper box. The generated mutations are shown in bold. (**B**) Mixtures of 1 μl of *in vitro* translated Pbx1 and 1 μl of En-2 or 0.5 μl of Hoxb-7, as indicated, were pre-incubated with the indicated micromolar amounts of peptide. Thirty minutes after addition of labeled probe 5′-GTCAATTAA-ATGATCAATCAATTTCG-3′ for En-2–Pbx or 5′-GTCAATTAAATGCGCATCAATCAATTTCG-3′ for Hoxb-7–Pbx, the products were analyzed by EMSA. (**C**) Mixtures of 1 μl of *in vitro* translated Pbx1 and 1 μl of En-2 or 0.5 μl of Hoxb-7 were incubated with the different peptides, as indicated, and analyzed by EMSA as in (B). (**D**) Mixtures of 1 μl of *in vitro* translated Pbx1 and 0.5 μl of Hoxb-7 were pre-incubated with different micromolar concentrations of the b-7, c-6 and b-3 peptides, as indicated, incubated with probe 5′-GTCAATTAAATGCGCATCAATCAATTTCG-3′ and analyzed by EMSA as in (B). (**E**) The quantities of the specific Hoxb-7–Pbx1 complexes of the EMSA of (D), as determined by PhosphorImager analysis (Molecular Dynamics, Sunnyvale, CA), are plotted as a function of the peptide concentration. Symbols are indicated in the figure. (**F**) Mixtures of 1 μl of *in vitro* translated Pbx1 and 0.5 μl of Hoxc-6 were pre-incubated with the indicated micromolar amounts of peptide. Thirty minutes after addition of labeled probe 5′-GTCAATTAAATGCGCATCAATCAATTTCG-3′, the products were analyzed by EMSA. (**G**) Quantitative representation of the results EMSA of (F), with on the *x*-axis the concentration of the used peptides and on the *y*-axis the radioactivity of the specific Hoxc-6–Pbx1 complex, as determined by PhosphorImager analysis. Symbols are indicated in the figure. The dark squares and bracket indicate specific complexes, fp indicates the free probe and the open diamonds mark complexes formed by unprogramed lysate.

the En-2–Pbx heterocomplex binds strongly, but the addition of En-2 wild-type peptide leads to competition for En-2–Pbx complexes (Figure 5B, lanes 1–5). To determine whether destabilization of the complex requires the conserved tryptophans, we incubated the mixture of En-2 and Pbx1 with the mutant En-2 peptide. Consistent with the previous results (Figure 2), this peptide could not destabilize the cooperative complex (Figure 5B, lanes 6–

9), indicating that the competition is specific for the wild-type EH2 domain. These data show that a peptide containing part of the EH2 domain is by itself capable of interaction with Pbx1.

To determine whether Hox proteins were similarly able to interact with Pbx, we synthesized a Hoxb-7 peptide spanning the hexapeptide (LAAESNFRIYPWMRSSGT-DRKRG, Figure 5A) and a Hoxb-7 mutant peptide in which a tryptophan has been mutated (LAAESNFRIYPY-MRSSGTDRKRG). These peptides were tested for competion of the Hoxb-7–Pbx1 complex. The Hoxb-7 peptide competes efficiently (Figure 5B, lanes 10–14), whereas the Hoxb-7 mutant peptide (Figure 5B, lanes 15–18) does not compete significantly.

To analyze whether the En-2 and Hox peptides can cross-compete the Hoxb-7–Pbx1 and En-2–Pbx1 complexes, respectively, mixtures of the appropriate *in vitro* translated proteins were incubated with increasing concentrations of En-2 or Hox peptides. The EMSA of the resulting complexes demonstrates that the En-2 wild-type peptide weakly, but significantly, competes for the Hoxb-7–Pbx1 complex, whereas its mutant peptide does not affect the cooperative complex at these concentrations (Figure 5C, lanes 1–5 and 6–9). The Hoxb-7 peptide shows substantial competition for the En-2 complex with Pbx, in contrast to its mutant counterpart (Figure 5C, lanes 10–14 and 15–18). These data suggest that EH2 and the Hox hexapeptide domain share, at least in part, an interaction surface on Pbx.

The sequences of the Hoxc-6 and Hoxb-3 hexapeptides are closely related to that of Hoxb-7 (Figure 5A), and both proteins can cooperate in DNA binding with Pbx. To determine whether the c-6 and b-3 peptides also have the ability to destabilize Hox–Pbx1 complexes, we incubated Hoxb-7–Pbx1 complexes with these peptides. The c-6 and b-3 peptides showed only weak competition of the Hoxb-7–Pbx complex at concentrations ranging from 5 to 20 μM. To determine if the weaker competition by the c-6 and b-3 peptides is due to a lower affinity for Pbx, mixtures of Pbx1 with either Hoxb-7 or Hoxc-6 proteins were incubated with larger amounts of c-6 or b-3 peptides. Both c-6 and b-3 peptides can reach the same level of competition as the b-7 peptide, but at 5- to 2.5-fold higher concentrations, respectively (Figure 5D–G). Thus, as expected, the hexapeptide regions of several different Hox proteins interact with Pbx, albeit with different affinities.

Hox peptides can stabilize intrinsic DNA binding activity of Pbx

Interestingly, the addition of Hoxb-7 peptide induces a new binding complex with faster mobility as compared with the Hoxb-7–Pbx1 heterodimer (Figure 5B and C). En-2 and both mutant peptides do not induce this complex (Figure 5B and C). The mobility of this complex is similar to that of Pbx1 complexed to DNA. EMSAs with Flag-tagged Pbx1 and anti-Flag antibodies indicated that these complexes indeed represent Pbx1 by itself bound to the probe (our unpublished data). To determine if this induction of DNA binding is seen in the absence of Hox or En-2 proteins, similar experiments with Pbx1 alone were performed. At a relatively low concentration of 5 μM, the b-7 peptide can stabilize the binding of Pbx1 monomers and homodimers to the DNA, and the amount of complexed

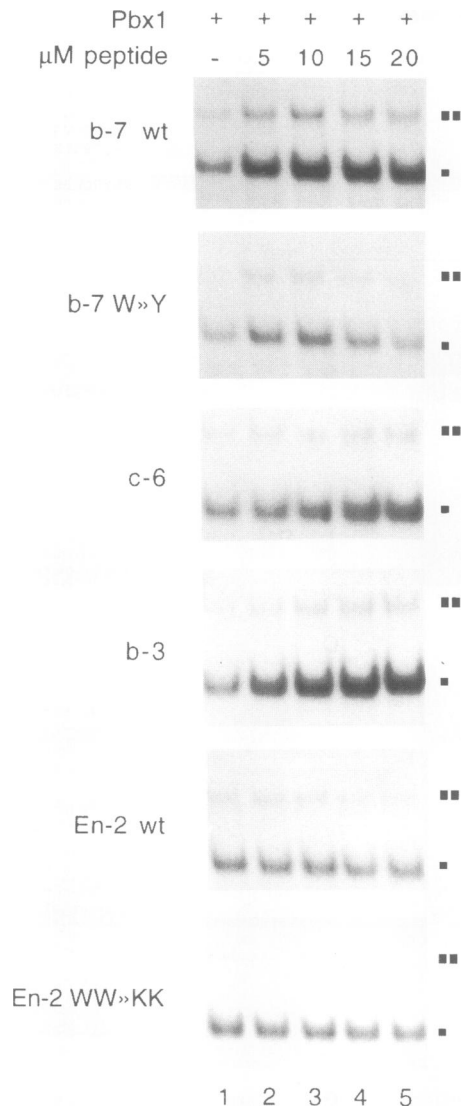

Fig. 6. Hox peptides, but not the En-2 peptide, can stabilize DNA binding of Pbx1. Each reaction contained 1 μl of *in vitro* translated Pbx1 incubated with increasing amounts of the b-7 wt, b-7 W>>Y, c-6, b-3, En-2 wt and En-2 WW>>KK peptides, as indicated on the left. Thirty minutes after addition of probe 5′-GTCAATTAAATGCGCAT-CAATCAATTTCG-3′, the products were analyzed by EMSA. The lower complex co-migrates with monomers of Pbx1, as indicated by a square, and the upper complex with Pbx1 homodimers bound to the probe, as indicated by a double square.

Pbx1 increases further with higher concentrations (Figure 6, upper panel). The b-7 peptide containing a tryptophan substitution was not capable of inducing Pbx binding activity, even at concentrations of 20 μM (Figure 6, second panel). Thus, the Hoxb-7 hexapeptide region by itself can modulate the DNA binding activity of Pbx1. In addition to b-7, both c-6 and b-3 peptides have the ability to enhance Pbx1 binding activity, albeit less effectively (Figure 6, third and fourth panel, and Figure 5D and F). Whereas the Hox peptides are all capable of inducing the binding activity of Pbx1, the En-2 peptide containing the tryptophan motif and its mutant counterpart do not have this property (Figure 6). Together, these results show that protein fragments comprising a Hox hexapeptide, but not the Engrailed interaction motif in EH2, are capable of enhancing the intrinsic DNA binding of Pbx1.

Discussion

A family of homeodomain proteins, containing the *Drosophila* gene product extradenticle and three mammalian proteins Pbx1, Pbx2 and Pbx3, has emerged recently. The Pbx class of proteins function as co-factors, which have little intrinsic DNA binding activity by themselves, but bind cooperatively to DNA with other families of homeodomain proteins. The other families include members of the *Drosophila* and vertebrate *hox* gene clusters (Chan *et al.*, 1994; van Dijk and Murre, 1994), an orphan homeodomain protein, designated STF-1 (Peers *et al.*, 1995) and members of the Engrailed family (van Dijk and Murre, 1994). The motif present in the Hox and orphan homeodomain proteins that is required for interaction with the Pbx members has been shown to contain a conserved hexapeptide (Chang *et al.*, 1995; Johnson *et al.*, 1995; Knoepfler and Kamps, 1995; Neuteboom *et al.*, 1995; Peers *et al.*, 1995; Phelan *et al.*, 1995). We have now extended those studies for Engrailed proteins. In Engrailed, the EH2 domain, shown to be conserved in all members and related but clearly distinct from the hexapeptide (Logan *et al.*, 1992), is required for cooperativity involving Pbx or extradenticle. Substitution of the conserved tryptophans in both murine and *Drosophila* Engrailed abolishes cooperativity. The amino acid sequence joining the EH2 domain to the homeodomain had been defined previously as the EH3 domain, again based on the extensive conservation of its residues (Logan *et al.*, 1992). Like the Hox and STF-1 linker peptide, EH3 does not tolerate gross deletions or insertions. We propose that the EH3 domain functions as a linker region similar to the linker peptide present in the Hox proteins and in the orphan homeodomain protein, STF-1 (Peers *et al.*, 1995).

Synthetic peptides, containing either the Hox or Engrailed Pbx interaction motifs, can destabilize both Hox–Pbx and Engrailed–Pbx cooperativity. Thus, both motifs have the ability to disrupt Pbx complexes that contain a heterologous partner. We speculate that Hox and Engrailed proteins interact with a common surface present in Pbx. Although peptides containing the Pbx interaction motif are capable of destabilizing heterologous complexes, it is evident that the peptides destabilize homologous cooperativity more efficiently than heterologous cooperativity. The most simple interpretation of this result is that the Engrailed and Hoxb-7 Pbx interaction motifs interact with a common structure present in Pbx1, but that additional residues present in Pbx interact specifically with either Engrailed or Hoxb-7. Which regions of Pbx are required for cooperative interaction is currently under investigation. From deletion and mutation analysis, it is clear that the requirements for Pbx1 cooperativity involving either Hox or Engrailed proteins are not identical. In summary, these data support a model in which Hox and Engrailed proteins share an interaction surface on the Pbx1 protein, but additional domains play a role as well.

The data also indicate that the various Hox hexapeptides have different affinities for Pbx. While the b-7 peptide destabilizes binding of both Hoxb-7–Pbx1 and Hoxc-6–Pbx1 heterodimers at relatively low concentrations, the c-6 peptide requires at least 5-fold more peptide to achieve 50% inhibition (Figure 5D–G). In addition, the b-7 peptide increased intrinsic Pbx1 DNA binding more as compared with the Hoxc-6 peptide (Figures 5D–G and 6). We would like to consider the possibility that the different relative affinities of the hexapeptides for Pbx1 may provide specificity to Hox–Pbx1 cooperativity. Different Hox proteins are expressed in various cell types during embryogenesis or lymphoid development. It is conceivable that among different possible combinations of Hox–Pbx complexes that can be generated certain combinations are favored based on the affinity of hexapeptide–Pbx interaction. Taken together with the relative concentrations of Hox proteins, the affinity of hexapeptide–Pbx interactions may ultimately decide which Hox–Pbx complexes are active within one given cell type.

It is intriguing that relatively low concentrations of the Hox peptides containing the Pbx interaction motif are sufficient to induce Pbx1 binding by itself. We have not detected specific DNA binding activity of the peptides by themselves, suggesting that a direct effect through neighboring DNA binding is unlikely. Hox peptides may simply induce a conformational change that releases an intramolecular inhibition of DNA binding by Pbx1. The question arises whether the peptides are capable of interacting with Pbx1 in solution in the absence of DNA. While these interactions may occur using very high concentrations of protein, we have not been able to detect Hox–Pbx interactions *in vitro* using nanomolar concentrations of protein. However, it is conceivable that the dissociation rates for these interactions are too high to be detected in these assays. Additional experimentation, for example using fluoresence polarization analysis, will be needed to address this issue further.

The EH2 domain is distinct from the hexapeptide present in Hox proteins with respect to the amount of conserved residues. However, there are some similarities. First, both contain conserved tryptophan residues that are required for cooperative DNA binding with Pbx. Second, the length of the linker region separating the Pbx interaction motif from the homeodomain in both Hox and Engrailed proteins is important for cooperativity. Third, both the Pbx interaction motifs of Hox and Engrailed cannot be grafted onto a heterologous DNA binding domain (Neuteboom *et al.*, 1995; Peers *et al.*, 1995; our unpublished data). This indicates that the N-terminal flanking regions and homeodomains are not interchangeable, consistent with the idea that the Pbx interaction domains in Hox and Engrailed proteins have evolved with their associated homeodomains.

Thus, Hox and Engrailed proteins each use a peptide located in close proximity to the homeodomain, allowing cooperative DNA binding to occur. Structural studies of the Antennapedia homeodomain protein have indicated that the region containing the hexapeptide and the linker does not form an ordered structure (Qian *et al.*, 1992; Billeter *et al.*, 1993). We have used circular dichroism spectroscopy to determine the structure of the Hoxb-7 peptide containing the Pbx interaction motif. The data show that the Hoxb-7 hexapeptide region is similarly unstructured in the absence of Pbx1 (our unpublished data). It will now be interesting to determine whether the Pbx interaction motifs in Hox and Engrailed proteins will adopt a structural change upon interaction with Pbx1. Such a structural change has been described for a domain

that allows two yeast homeodomain proteins, MATa1 and MATα2, to bind cooperatively to DNA (Li *et al.*, 1995).

There is a strong parallel between the interactions involving selector gene products and Pbx, and the two yeast homeodomain proteins a1 and α2 (Johnson, 1992). α2 interacts with two proteins, a1 and MCM1, using flexible unstructured peptides that are located either C- or N-terminal of the homeodomain (Vershon and Johnson, 1993; Stark and Johnson, 1994). Similarly, flexible unstructured peptides present N-terminal of the Hox and Engrailed homeodomains interact with exd and Pbx. Thus, exd and Pbx have evolved as factors capable of interacting with distinct peptides present in homeodomain-containing proteins. It may very well be that the presence of peptides flanking the homeodomain in selector proteins is an ancient feature of homeodomain proteins, allowing them to modulate the DNA binding affinity and specificity of their partners.

Materials and methods

Plasmids

All constructs for *in vitro* transcription–translation were derivatives of the vector pSP64 (Promega, Madison, WI), the modified forms pSP64-ATG (van Dijk and Murre, 1994) or its derivative containing three out-of-frame stop codons, pSS (Neuteboom *et al.*, 1995). To make the murine En-2 deletion constructs, PCR fragments spanning EH2–4 or EH4 itself were cloned into the *Eco*RV site of pSS. The En-2 mutant forms W1>>K, W4>>K and WW>>KK were generated by two-step PCR of pSP64/En-2, containing the full-length En-2 coding sequence (van Dijk *et al.*, 1995) with primers bearing the mutation. The *Drosophila* en coding sequence was mutated by two-step PCR of pSP64-ATG/en, that contains the entire coding region of the *engrailed* gene (van Dijk and Murre, 1994). The En-2 del 6, del 11 and plus 10 constructs were all derived from pSS/EH2–4 by using two-step PCR with oligonucleotides overlapping the desired junctions (see Figure 4A). To equip the Pbx coding regions with an N-terminal Flag tag, the oligonucleotide 5′-AAGCTTCAGCATGGACTACAAGGACGACGATGACAAGGGATC-CATGG-3′ was ligated in the *Hin*dIII and *Nco*I sites of pSP64-ATG, resulting in pSP64Flag. Subsequently, complete coding regions of Pbx1, Pbx2 and Pbx3 were cloned in-frame with the underlined Flag start codon. The untagged Pbx1 translation construct was a modified form of pSP65-Pbx1, having the 5′ end of the pSP64-ATG polylinker and its start codon modified to an *Nco*I site. The exd product was translated from pSP64-ATG/exd (van Dijk and Murre, 1994). The Hoxb-7 translation construct was described in van Dijk *et al.* (1995). The Hoxc-6 full-length coding region was amplified from a cDNA obtained from E.Boncinelli (Magli *et al.*, 1991) and cloned in pSP64Flag. The integrity of all cloned amplification products was verified by DNA sequencing.

In vitro translation and electrophoretic mobility shift assays (EMSA)

All proteins used in these experiments were generated using the Promega SP6 TNT rabbit reticulocyte lysate-coupled transcription–translation system according to the manufacturer's protocol (Promega, Madison, WI). EMSAs were done as described (van Dijk and Murre, 1994). One to 2 μl of each reticulocyte lysate-translated protein, to a maximum of 4 μl total, was used per reaction. All reactions were performed at room temperature and adjusted to contain equal amounts of reticulocyte lysate. Thirty minute incubations with 1 μl of the anti-Flag antibody (Eastman Kodak, New Haven, CT) were done after pre-incubation of translated proteins with DNA probe for 30 min. The ^{32}P end-labeled DNA probes used were 5′-GTCAATTAAATGATCAATCAATTTCG-3′ for Engrailed–Pbx cooperative binding and 5′-GTCAATTAAATGCGCAT-CAATCAATTTCG-3′, for Hox–Pbx cooperativity.

Peptide synthesis and competition

Peptides were synthesized (Chiron Mimotopes, San Diego, CA) and purified by reverse-phase HPLC. The amino acid sequences are as follows: En-2 wt, TLGAQPMLWPAWVYCTRYSDRPS; En-2 WW>> KK, TLGAQPMLKPAKVYCTRYSDRPS; Hoxb-7 wt, LAAESNFR-IYPWMRSSGTDRKRG; Hoxb-7 W>>Y, LAAESNFRIYPYMRSSG-TDRKRG, Hoxc-6, QDQKASIQIYPWMQRMNSHSGVG; Hoxb-3, TNSTLTKQIFPWMKESRQTSKLK. Peptides were dissolved in water and the pH was adjusted by the addition of 1 M HEPES–KOH, pH 7.4. Concentrations were determined spectrophotometrically (Gill and von Hippel, 1989) and appropriate dilutions were made in 20 mM HEPES, pH 7.4/0.1 mg/ml bovine serum albumin. Peptides in concentrations ranging from 5 to 100 μM and translated proteins were pre-incubated for at least 5 min before labeled DNA probe was added.

Acknowledgements

We would like to thank Edoardo Boncinelli, Alexandra Joyner, Saskia Neuteboom and John McWhirther for the gift of several plasmids, and Jaume Canaves for help with the peptide purification. This work was sponsored by the National Institutes of Health, the Council for Tobacco Research and the Edward Mallinckrodt Jr Foundation. L.T.C.P. was supported by a fellowship of the Dutch Cancer Society (KWF).

References

Billeter,M., Qian,Y.Q., Otting,G., Muller,M., Gehring,W. and Wüthrich,K. (1993) Determination of the nuclear magnetic resonance solution structure of an Antennapedia homeodomain–DNA complex. *J. Mol. Biol.*, **234**, 1084–1093.

Bürglin,T.R. (1994) A comprehensive classification of homeobox genes. In Duboule,D. (ed.), *Guidebook to the Homeobox Genes*. Oxford University Press, Oxford, UK, pp. 27–71.

Chan,S.-K. and Mann,R.S. (1993) The segment identity functions of Ultrabithorax are contained within its homeo domain and carboxy-terminal sequences. *Genes Dev.*, **7**, 796–811.

Chan,S.-K., Jaffe,L., Capovilla,M., Botas,J. and Mann,R.S. (1994) The DNA binding specificity of ultrabithorax is modulated by cooperative interactions with extradenticle, another homeodomain protein. *Cell*, **78**, 603–615.

Chang,C.-P., Shen,W.-F., Rozenfeld,S., Lawrence,H.J., Largman,C. and Cleary,M.L. (1995) Pbx proteins display hexapeptide-dependent cooperative DNA binding with a subset of Hox proteins. *Genes Dev.*, **9**, 663–674.

Desplan,C., Theis,J. and O'Farrell,P.H. (1988) The sequence specificity of homeodomain–DNA interaction. *Cell*, **54**, 1081–1090.

Duboule,D. and Dollé,P. (1989) The structural and functional organization of the murine *HOX* gene family resembles that of *Drosophila* homeotic genes. *EMBO J.*, **8**, 1497–1505.

Flegel,W.A., Singson,A.W., Margolis,J.S., Bang,A.G., Posakony,J.W. and Murre,C. (1993) *Dpbx*, a new homeobox gene closely related to the human proto-oncogene *pbx1*: molecular structure and developmental expression. *Mech. Dev.*, **41**, 155–161.

Gibson,G., Schier,A., Lemotte,P. and Gehring,W.J. (1990) The specificities of sex combs reduced and antennapedia are defined by a distinct portion of each protein that includes the homeodomain. *Cell*, **62**, 1087–1103.

Gill,S.C. and von Hippel,P.H. (1989) Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.*, **182**, 319–326.

Hoey,T. and Levine,M. (1988) Divergent homeo box proteins recognize similar DNA sequences in *Drosophila*. *Nature*, **332**, 858–861.

Johnson,A. (1992) A combinatorial regulatory circuit in budding yeast. In McKnight,S.L. and Yamamoto,K.R. (eds), *Transcriptional Regulation*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 975–1006.

Johnson,F.B., Parker,E. and Krasnow,M.A. (1995) extradenticle protein is a selective cofactor for the *Drosophila* homeotics: role of the homeodomain and YPWM amino acid motif in the interaction. *Proc. Natl Acad. Sci. USA*, **92**, 739–743.

Kamps,M.P., Murre,C., Sun,X.-H. and Baltimore,D. (1990) A new homeobox gene contributes the DNA binding domain of the t(1;19) translocation protein in pre-B ALL. *Cell*, **60**, 547–555.

Knoepfler,P.S. and Kamps,M.P. (1995) The pentapeptide motif of Hox proteins is required for cooperative DNA binding with Pbx1, physically contacts Pbx1, and enhances DNA binding by Pbx1. *Mol. Cell. Biol.*, **15**, 5811–5819.

Krumlauf,R. (1994) Hox genes in vertebrate development. *Cell*, **78**, 191–201.

Kuziora,M.A. and McGinnis,W. (1989) A homeodomain substitution changes the regulatory specificity of the deformed protein in *Drosophila* embryos. *Cell*, **59**, 563–571.

Laughon,A. and Scott,M. (1984) Sequence of a *Drosophila* segmentation gene: protein structure homology with DNA binding proteins. *Nature*, **310**, 25–31.

Lawrence,P.A. and Morata,G. (1994) Homeobox genes: their function in *Drosophila* segmentation and pattern formation. *Cell*, **78**, 181–189.

Li,T., Stark,M.R., Johnson,A.D. and Wolberger,C. (1995) Crystal structure of the MATa1/MATα2 homeodomain heterodimer bound to DNA. *Science*, **270**, 262–269.

Lin,L. and McGinnis,W. (1992) Mapping functional specificity in the Dfd and Ubx homeo domains. *Genes Dev.*, **6**, 1071–1018.

Logan,C., Hanks,M., Nobletopham,S., Nallainathan,D., Provart,N. and Joyner,A. (1992) Cloning and sequence comparison of the mouse, human, and chicken engrailed genes reveal potential functional domains and regulatory regions. *Dev. Genet.*, **13**, 345–358.

Lu,Q., Knoepfler,P.S., Scheele,J., Wright,D.D. and Kamps,M.P. (1995) Both Pbx1 and E2A-Pbx1 bind the DNA motif ATCAATCAA cooperatively with the products of multiple murine Hox genes, some of which are oncogenes themselves. *Mol. Cell. Biol.*, **15**, 3786–3795.

Magli,M.C., Barba,P., Celetti,A., De Vita,G., Cillo,C. and Boncinelli,E. (1991) Coordinate regulation of *HOX* genes in human hematopoietic cells. *Proc. Natl Acad. Sci. USA*, **88**, 6348–6352.

McGinnis,W. and Krumlauf,R. (1992) Homeobox genes and axial patterning. *Cell*, **68**, 283–302.

Monica,K., Galili,N., Nourse,J., Saltman,D. and Cleary,M.L. (1991) *PBX2* and *PBX3*, new homeobox genes with extensive homology to the human proto-oncogene *PBX1*. *Mol. Cell. Biol.*, **11**, 6149–6157.

Neuteboom,S.T.C., Peltenburg,L.T.C., van Dijk,M.A. and Murre,C. (1995) The hexapeptide LFPWMR in Hoxb-8 is required for cooperative DNA binding with Pbx1 and Pbx2 proteins. *Proc. Natl Acad. Sci. USA*, **92**, 9166–9170.

Nourse,J., Mellentin,J.D., Galili,N., Wilkinson,J., Stanbridge,E., Smith,S.D. and Cleary,M.L. (1990) Chromosomal translocation t(1; 19) results in synthesis of a homeobox fusion mRNA that codes for a potential chimeric transcription factor. *Cell*, **60**, 535–545.

Peers,B., Sharma,S., Johnson,T., Kamps,M.P. and Montminy,M. (1995) The pancreatic islet factor STF-1 binds cooperatively with Pbx to a regulatory element in the somatostatin promoter: importance of the FPWMK motif and of the homeodomain. *Mol. Cell. Biol.*, **15**, 7091–7097.

Peifer,M. and Wieschaus,E. (1990) Mutations in the *Drosophila* gene extradenticle affect the way specific homeo domain proteins regulate segmental identitity. *Genes. Dev.*, **4**, 1209–1223.

Phelan,M.L., Rambaldi,I. and Featherstone,M.S. (1995) Cooperative interactions between Hox and Pbx proteins mediated by a conserved peptide motif. *Mol. Cell. Biol.*, **15**, 3989–3997.

Pöpperl,H., Bienz,M., Studer,M., Chan,S.-K., Aparicio,S., Brenner,S., Mann,R.S. and Krumlauf,R. (1995) Segmental expression of Hoxb-1 is controlled by a highly conserved autoregulatory loop dependent upon exd/pbx. *Cell*, **81**, 1031–1042.

Qian,Y.Q., Otting,G., Furukubo-Tokunaga,K., Affolter,M., Gehring,W.J. and Wüthrich,K. (1992) NMR structure determination reveals that the homeodomain is connected through a flexible linker to the main body in the *Drosophila* Antennapedia protein. *Proc. Natl Acad. Sci. USA*, **89**, 10738–10742.

Rauskolb,C. and Wieschaus,E. (1994) Coordinate regulation of downstream genes by extradenticle and the homeotic selector genes. *EMBO J.*, **13**, 3561–3569.

Rauskolb,C., Peifer,M. and Wieschaus,E. (1993) Extradenticle, a regulator of homeotic gene activity, is a homolog of the homeobox-containing human proto-oncogene *pbx1*. *Cell*, **74**, 1101–1112.

Stark,M.R. and Johnson,A.D. (1994) Interaction between two homeodomain proteins is specified by a short C-terminal tail. *Nature*, **371**, 429–432.

van Dijk,M.A. and Murre,C. (1994) extradenticle raises the DNA binding specificity of homeotic selector gene products. *Cell*, **78**, 617–624.

van Dijk,M.A., Peltenburg,L.T.C. and Murre,C. (1995) Hox gene products modulate the DNA binding activity of Pbx1 and Pbx2. *Mech. Dev.*, **52**, 99–108.

Vershon,A. and Johnson,A. (1993) A short, disordered protein region mediates interactions between the homeodomain of the yeast α-2 protein and the mcm1 protein. *Cell*, **72**, 105–112.

Wilkinson,D.G., Bhatt,S., Cook,M., Boncinelli,E. and Krumlauf,R. (1989) Segmental expression of Hox-2 homeodomain-containing genes in the developing mouse hindbrain. *Nature*, **341**, 405–409.

Zeng,W., Andrew,D., Mathies,L., Horner,M. and Scott,M. (1993) Ectopic expression and function of the *Antp* and *Scr* homeotic genes—the N-terminus of the homeodomain is critical to functional specificity. *Development*, **118**, 339–352.

Received on January 19, 1996; revised on March 14, 1996